Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 65 (2015) 744 – 753

Procedia Computer Science

International Conference on Communication, Management and Information Technology (ICC 2015)

E-inclusion modeling for blended e-learning course

Ieva Vitolina*

Riga Technical University, Kalku street 1, Riga, LV-1658, Latvia

Abstract

This study addresses the e-inclusion problem that relates to the inclusion of as many individuals as possib enjoy benefits of information and communication technology. Despite the fact that European Union accepte inclusion declaration in 2006 which aims to reduce disparities that exist among individuals and to improve the l of e-skills among people, nowadays e-inclusion problem still exists. Therefore it is necessary to find out approach to promote e-inclusion in society. We propose a more nuanced design approach that takes into acc student's satisfaction with e-learning environment and e-materials, student's ability to learn, instructor willingne share knowledge and others factors. Moreover we believe that e-inclusion means not only high level of digital s but also the usage of these digital skills to benefit from technologies. To obtain predictors for algorithms we d inclusion data domain study based on knowledge management theory. The aim of proposed work is to prese inclusion theoretical model which is based on integration of several algorithms as multiply linear regression cluster analysis. These algorithms were calculated based on statistical data obtained on evaluating a group of hundred blended e-course learners. In this paper we propose architecture designed to predict e-inclusion degr student based on machine learning and intelligent agent approach. We identified two main processes in th inclusion prediction system. The first process consists of agent learning activities. Intelligent agents learn the r appropriate algorithm to predict e-inclusion degree of student based on linear regression or cluster analysis. second process includes activities to predict e-inclusion degree of student. This process covers analysis c inclusion risks and communication between student and instructor also. Proposed e-inclusion model consists of diagram, use cases diagrams and main algorithms of the system. As the result of the e-inclusion model is predic of e-inclusion degree of person as well as e-inclusion risk factors for person, for instance inappropriate e-lear materials or no interest to learn, or dissatisfaction with e-learning environment, or others factors.

* Corresponding author. Tel.: +371 29215073.
E-mail address: ieva.vitolina@rtu.lv

Peer-review under responsibility of Universal Society for Applied Research
doi:10.1016/j.procs.2015.09.019

Peer-review under responsibility of Universal Society for Applied Research

Keywords: E-inclusion; e-learning, machine learning

1. Introduction

This study aims to address the e-inclusion problem that was outlined in the EU Digital Agenda 2020 that r to the inclusion of as many individuals as possible to enjoy the benefits of information and communic technology (ICT)[1]. Nowadays the digital divide goes beyond the issue of access to technology. Focus has sl from access to ICT to digital skills and the meaningful use of ICT[2]. There is a gap between knowing to d practical usage of digital skills. Learning a new skill and using it are two separate steps[3]. The 2010 OECD r stated that a second digital divide separates those with the competencies and skills that benefits from compute from those without these advantages[4].

Several studies indicate that there is a need to look for factors that characterize the e-included individ Therefore, it is necessary to identify the factors that influence e-inclusion process so that individuals learn technologies and use them meaningfully. There are currently no comprehensive methods to monitor meaningfu of digital skills in order to prevent the ICT usage gap.

There is no special technology for e-inclusion prediction. Usually systems predict whether students drop c complete course. Machine learning approach is used for student achievement and other event prediction. Ma learning and agent technologies are integrated with a particular interest on applying agent-based solutio supervised learning[6].

This study contributes to research of the meaningful ICT use in blended learning context. In this pape propose a new system architecture designed to predict e-inclusion degree of student based on machine learnin intelligent agent approach.

The paper is organized as follows. In the section 2 technologies and methods of prediction are described. section 3 introduces e-inclusion model design. In the section 4 e-inclusion prediction algorithms are presented section 5 is conclusion part.

2. Technologies and methods of the e-inclusion prediction

2.1. Review of literature

According to Strickland[7] predictive analytics is an area of data science that deals with extracting inform from data and using it to predict trends and behavior patterns. A predictive model is a statistical model or ma learning model used to predict future behavior based on past behavior. Machine learning approach is used into oriented systems to train agents. In the literature machine learning and agent oriented system development m has following steps: agent selection, problem domain analysis and data selection and pre-processing, selecti machine learning method and algorithm, model evaluation and implementation with prediction function[8].

Machine learning-based system has two main processes: (1) system training process and (2) the pred process (Fig.1). Each of these two phases is subdivided into several steps and presented in the Figure 1.

Machine learning process begins with analysis of problem domain and data selection[9]. More appro selection of data is possible after deep problem domain analysis.

Data pre-processing aims to create a sample database containing both the training and test data for the model

Next step is selection of algorithm and method for prediction. Training data are transferred to the tra algorithm. At this stage a predictive model is build. In the literature this stage is known as creation of the know base[11]. It should be noted that no machine learning method or algorithm is clearly better than another, bu machine learning method should be assessed with test data sets for more precise results. Data analysis professi as the most commonly used data mining algorithms present linear regression, decision trees and cluster analysis[1

Model accuracy evaluation is very important step. The goal of evaluation stage is to make sure that the predi model is able to predict with other data as accurately as with training data set[13]. At this stage, model valid method is used for the test data set. If the model satisfies the requirements of precision criteria, then begins the stage - the predicting process. If the model does not meet the criteria of accuracy, the system begins learning pro again.

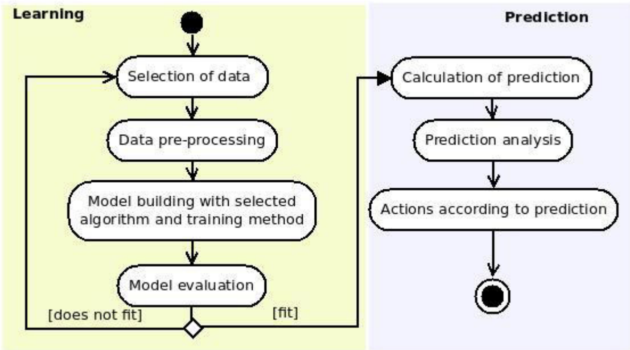


Fig. 1. The main processes of the machine learning-based predicting system.

Prediction stage begins with receiving data and calculation of prediction[14]. At this stage the system uses lea knowledge to perform calculations[9, 11].

The next stage is analysis and interpretation of results[13]. For instance, at this stage, it is determined whether t is a risk and what factors predict the risk.

The final stage is the system action according to the obtained prediction[10]. Depending on the tasks this p differs for various systems. For instance, if the prognosis is favorable to the student, the system can analyze the student. If results point to the risks, the system can inform the student, instructor, educational institution representative may initiate risk prevention activities.

2.2. E-inclusion domain analysis and selection of algorithms for prediction

Following to the schema given above author analyzed e-inclusion domain and determined most approp methods and algorithms for predicting e-inclusion.

Problem domain analysis and data selection. The e-inclusion domain research employed Enterprise Knowl Development (EKD) framework[15]. Author categorized the major obstacles to overcoming the ICT gap as well a most significant factors that promote e-inclusion. Author concluded that meaningful use of ICT is basic factor w promotes e-inclusion for individuals[16, 17].

Author's previous research results indicated that essential data for prediction are following: the quality learning materials and environment, instructor's support, student willingness and ability to learn. These data attributes that could impact learning carry-over and promoted e-inclusion[18]. This previous study was base evaluating a group of five hundred vocational teachers who were the learners in the blended e-learning co "Improvement of ICT skills". Based on vocational teachers inquiry the blended e-learning course contained important topics for teachers to develop digital competence. The topics of the course related to the improveme instrumental knowledge and skills for tool and media usage, advanced skills and knowledge for communicat information management, and continued learning and meaningful participation in a knowledge society. We anal ten topics: Setup of peripherals, Image scanning, Web page design, PDF files, Computer security, MS Access, V processing, E-learning materials, Social networks, and E-mail. Each topic included theoretical material in video text format and tests for knowledge assessment. In addition, we assigned practical exercise to apply the knowl gained.

Selection of algorithms. We concluded that linear regression and cluster analysis are appropriate methods
inclusion prediction.

We identified that cluster analysis with WEKA software are appropriate method for e-inclusion prediction[1
used EM (Expectation-Maximization) and KMeans as clustering algorithms[20]. We observed that values of attri
are different if e-inclusion level of student changes. The main attributes of clasters were[21]:

- student's interest in learning;
- student's ability to learn;
- instructor's willingness to share knowledge;
- student's assessment of e-learning environment;
- student's evaluation of e-learning materials;
- student's knowledge level before learning;
- student's digital skill level;
- student's predicted use of newly learned skills.

We did calculations of multiple regressions and recommended algorithms to determine the relationship bet
several instructor and student variables to design e-inclusion model[22, 23].

In context of knowledge transfer in the individual level student's e-inclusion degree is possible to predict
linear regression model, where predictors are following independent variables:

- student's interest in learning;
- the student's ability to learn;
- instructor's willingness to share knowledge;
- student's assessment of e-learning environment;
- student's evaluation of e-learning materials.

Also we developed a formula for linear regression equations for e-inclusion degree in context of knowledge
acceleration. As potential predictor for effective delivery of different topics of an e-learning course we obt
knowledge flow acceleration. The results indicated that one of the factors for determining digital skills pra
usage probability in the e-inclusion model for an e-learning course was related to knowledge flow acceleration.

We concluded that e-inclusion linear regression models are different for different e-learning topics. It i
possible to describe all of the course topics with a single linear regression equation.

2.3. Comparison with other method

There are no widely distributed studies which investigate in machine learning context connection of digital
learning process and digital skills practical usage after e-course completing. Traditionally literature relates h
avoid student's drop outs. E-inclusion means not only to obtain digital skills but also to use them. In our stud
are interested in prediction of practical usage of newly learned digital skills. Traditional dropping-out pred
models are based on demographic and financial attributes about students[24]. We selected data based on know
management theory by Nissen[25], for instance instructor's ability to share knowledge, student's satisfaction w
learning environment and e-learning materials, student's ability to learn and knowledge flow acceleration. Us
different supervised algorithms or unsupervised algorithms have been applied in the prediction of student drop
out systems. Linear regression, Bayesian networks, decision trees for the deduction of rules are used rega
student learning results prediction[26, 27, 28]. Prediction of student withdrawal prior to completion of their degr
based on unsupervised algorithm - Self-Organizing Map, clustering[24, 26]. To provide higher accuracy of pred
we used ensemble technique to combine both supervised and unsupervised approach.

3. E-inclusion model design

3.1. Methodology

E-inclusion model building is based on machine learning approach and MASITS methodology. MASITS is a life cycle methodology for agent based system development[29,30]. Similar to the conventional information sys development process MASITS consists of the following phases: analysis, design (divided into two stages: exte and internal design), implementation, testing, deployment and maintenance. This paper presents an e-inclu model analysis phase that includes two consecutive steps. The first step is the goal modelling resulting in the diagram. At the second step the use case model is created.

The model contains agents, which learn algorithms for predicting and tend towards the most accurate predic Different notations of the Unified Modeling Language (UML) are used to present the e-inclusion model[31].

3.2. Goals model of the e-inclusion prediction system

Figure 2 presents goals diagram for e-inclusion model. Development of goals is based on machine lear approach. There are two main processes: system training and predicting.

The e-inclusion goals modeling is based on results of previous research where author concluded that

- e-inclusion degree of student is possible to predict by linear regression and cluster analysis methods;
- system must build predictive models for each topic separately based on sample data; the predictive mo differs for various topics;
- due to the fact that the student data changes in the real time the system must learn dynamic data to incr accuracy of the predictive model.

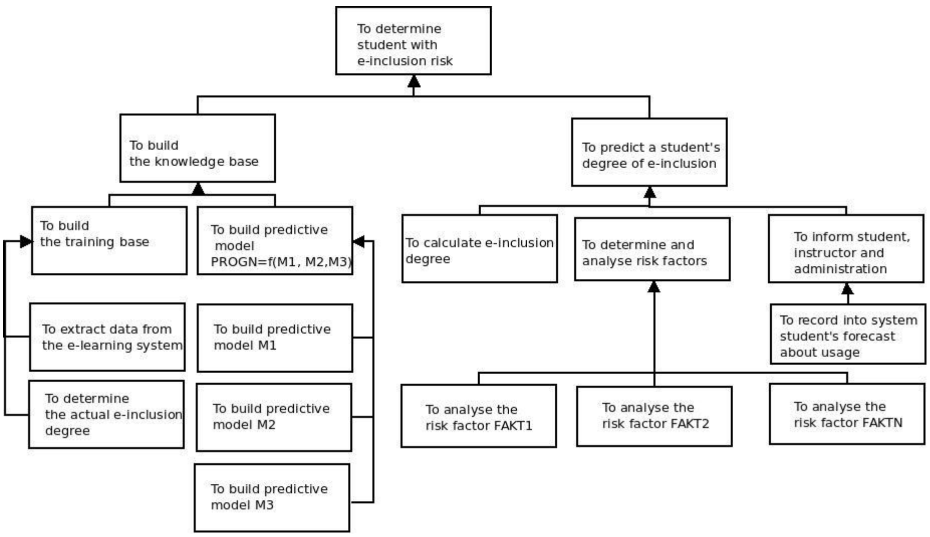


Fig.2. Goals model for e-inclusion prediction system.

The main goal of the e-inclusion model is to determine student with e-inclusion risk. Sub goals are to buil knowledge base and to predict a student's degree of e-inclusion. The building of the knowledge base includes

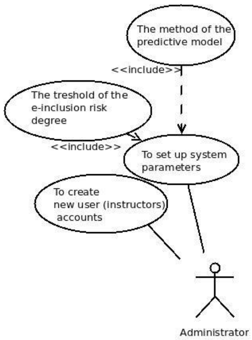


Fig. 3. Administrator use case of e-inclusion model.

following tasks: to extract student data from e-learning system and to supplement these data with actual e-incl degree of previous tested students. The building of knowledge base consists of finding the e-inclusion pred algorithm PROGN=f(M1, M2, M3) where M1 and M2 are models based on linear regression, M3 is based on c analysis. While goal of prediction a student's degree of e-inclusion has following tasks: to calculate e-incl degree with algorithm PROGN=f(M1, M2, M3), to determine whether there is risk factors related to e-lea environment, e-learning materials, instructor's willingness to share knowledge, to analyze these risk factors a inform student, instructor and administrator to avoid exclusion.

3.3. Main use cases of the e-inclusion prediction model

At the beginning of the use case modeling the actors are defined. The main actors of the e-inclusion mod instructor, system administrator and interface module which aims to initiates data acquisition from e-lea systems for further analysis and the student's e-inclusion degree determination.

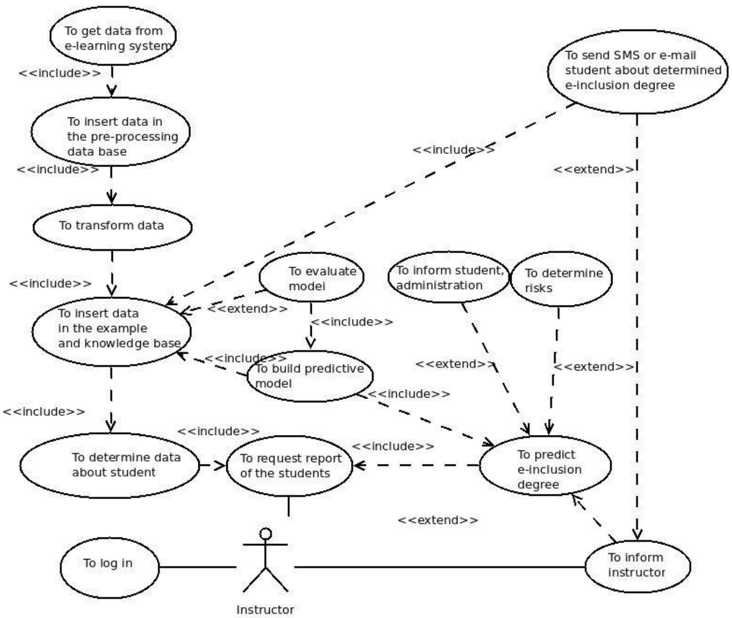


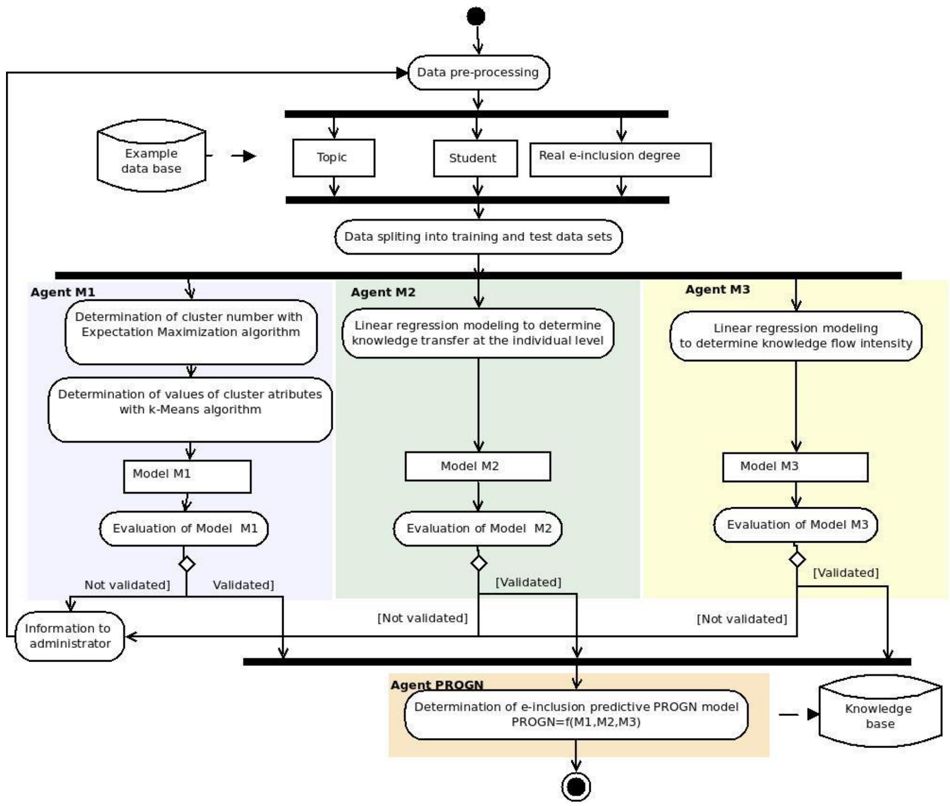
Fig. 4. Instructor use case of e-inclusion model

System administrator creates accounts for instructors, set up system parameters, to determine threshold (inclusion degree etc. (Fig. 3).

Use case of instructor is presented in the figure 4. After log in the system instructor requires a statement o student e-inclusion degree for specific learning topic. The forecasting process is initiated. If the forecast indi that the student has exclusion risk then initiated a risk analysis process to determine which risk factors affec student's level of e-inclusion. The instructor receives information about student and relevant risk factor. To be ab create a predictive model a sample data base was created by transforming the e-learning system data. The sa data base is updated with the student's forecast about digital skills usage and the actual degree of student inclusion

4. Algorithms of the e-inclusion prediction model

Figure 5 and figure 6 present learning and prediction algorithms in UML Activity diagram. Learning phase s with the data pre-processing and insertion in the sample base. Following data is necessary to include in the sa database: topic, data about student and it behaviour that indicates a student's actual level of e-inclusion. Sample base is split into training data and test data sets. Next, the system performs the learning process with training The learning process uses agent oriented approach, each predictive model is created by agent, whose task is to the best predictive model for appropriate method - linear regression or cluster analysis. In the next step three m are evaluated and the model with the best accuracy is selected. At this stage intellectual agent approach is u



Data obtained are included in the knowledge database and used for the forecasting process next.

Fig. 5. Learning algorithm of e-inclusion model.

Prediction algorithm use following input data: e-course topic and student's data. If the prediction show: there is a risk of exclusion then system identifies which risk factors are included in the model PROGN(M1, M3), for instance e-learning environment, learning materials, or instructor's ability to share knowledge. Bas the obtained risk factors the instructor decides further action. Instructor can changes something in communic with the student, or offer other training materials, etc. 8 to 10 weeks after e-learning course completing the sy sends the student SMS or e-mail with question about the practical use of digital skills. The student's answ recorded in the sample database.

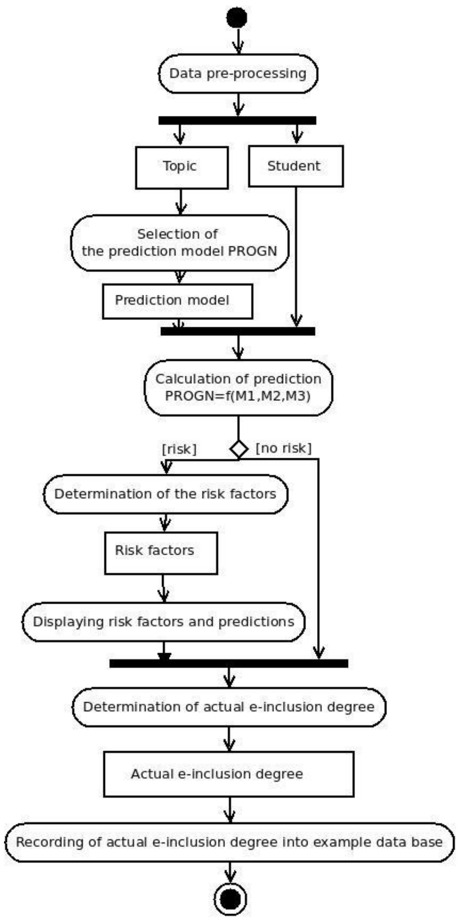


Fig. 6. Prediction algorithm of e-inclusion model.

5. Conclusions

This study contributes to research of e-inclusion in context of blended e-learning. In this paper we pr architecture designed to predict e-inclusion degree of student based on machine learning and intelligent approach. We identified two main processes in the e-inclusion prediction system. The first process consis learning activities. Intelligent agents learn the most appropriate algorithm to predict e-inclusion degree of stu based on linear regression or cluster analysis. The second process include activities to predict e-inclusion degr student. This process cover analysis of e-inclusion risks and communication between student and instructor Proposed e-inclusion model consists of goal diagram, use cases diagrams and main algorithms of system.

Future work is to develop a prototype of e-inclusion model based on our proposed architecture. Also in our fu work we will evaluate effectiveness of this e-inclusion prediction system. Our purpose is to evaluate prototyp the group of instructors. The prototype will be evaluated twice: by providing system demonstration sessio instructors and by performing the survey. The proposals of instructors will be taken into account in the fu development of a prototype. In our future work we will focus on the application of the proposed e-inclusion m in the Distance Education Study Centre of Riga Technical University.

Acknowledgements

This research has been supported by a grant from the European Regional Development Fund (ERFD/ER project "New User behavioral interpretation algorithms to facilitate an efficient transfer of knowledge within a ecosystem (JAUZI)", Grant Agreement 2013/0071/2DP/2.1.1.1.0/13/APIA/VIAA/023.

References

1. European Commission. (2012). Digital Agenda Scoreboard 2011, Pillar 6: Digital Competence in the Digital Agenda https://ec.europa.eu/digital-agenda/sites/digital-agenda/files/digitalliteracy.pdf
2. Hargittai, E., & Hinnant, A. (2008). Digital Inequality: Differences in Young Adults' Use of the Internet. Communication Research, 602-621
3. Lerchner, Alexander, Giancarlo La Camera, and Barry Richmond. "Knowing without doing." Nature neuroscience 10.1 (2007): 15-17.
4. OECD, DAC. "Development Cooperation Report 2010." (2010): 26.
5. De Haan, Jos. "A multifaceted dynamic model of the digital divide." It & Society 1.7 (2004): 66-88.
6. Jędrzejowicz, P. (2011). Machine learning and agents. In Agent and Multi-Agent Systems: Technologies and Applications (pp. 2-15). Sp Berlin Heidelberg.
7. Strickland, J. Predictive Analytics using R.
8. Sardinha, J. A. R. P., et al. "The agent learning pattern." Fourth Latin American Conference on Pattern Languages of Programming, SugarLoafPLoP. Vol. 4. 2004.
9. Ning Chen; Hoi, S.C.H.; Xiaokui Xiao, "Software process evaluation: A machine learning approach," Automated Software Engineering (. 2011 26th IEEE/ACM International Conference on , vol., no., pp.333,342, 6-10 Nov. 2011
10. Ribeiro De Carvalho Martinho, V.; Nunes, C.; Minussi, C.R., "An Intelligent System for Prediction of School Dropout Risk Group in F Education Classroom Based on Artificial Neural Networks," Tools with Artificial Intelligence (ICTAI), 2013 IEEE 25th Interna Conference on , vol., no., pp.159,166, 4-6 Nov. 2013
11. Wang Hua; Ma Cuiqin; Zhou Lijuan, "A Brief Review of Machine Learning and Its Application," Information Engineering and Con Science, 2009. ICIECS 2009. International Conference on , vol., no., pp.1,4, 19-20 Dec. 2009, doi: 10.1109/ICIECS.2009.5362936
12. Rexer Analytics. (2013). Data Miner Survey. http://www.rexeranalytics.com/Data-Miner-Survey-Results-2013.html
13. Halkidi, M., Batistakis, Y., & Vazirgiannis, M. (2001). On clustering validation techniques. Journal of Intelligent Information Systems, 3), 107-145.
14. Petkovic, D.; Okada, K.; Sosnick, M.; Iyer, A.; Shenhaochen Zhu; Todtenhoefer, R.; Shihong Huang, "Work in progress: A machine lea approach for assessment and prediction of teamwork effectiveness in software engineering education," Frontiers in Education Confe (FIE), 2012 , vol., no., pp.1,3, 3-6 Oct. 2012
15. Bubenko, J. A., Anne Persson, and J. Stirna. "User guide of the knowledge management approach using enterprise knowledge patt Stockholm (Sweden), Department of Computer and Systems Science, Royal Institute of Technology (2001).
16. Vītoliņa I. A Knowledge Management Model to Promote E-Inclusion // 7th International JTEFS/BBCC Conference "Susta Development. Culture. Education" : Research and Implementation of Education for Sustainable Development: Proceedings, La Daugavpils, 5.-8. maijs, 2009. - 6-38. lpp.
17. Vītoliņa I. E-inclusion modeling to improve digital skills of society. Liepājas Universitātes 13. starptautiskā zinātniskā konfer "Sabiedrība un kultūra" rakstu krājums, Liepāja, 2011. 869. –878. lpp.
18. Vītoliņa I., Kapenieks A. 2012. A Study of the e-inclusion process in a real-life e-course delivery context. 10th International JTEFS/E Conference "Sustainable Development. Culture. Education": May 22 – 25, 2012 Savonlinna, Finland.
19. Hall, Mark, et al. "The WEKA data mining software: an update." ACM SIGKDD explorations newsletter 11.1 (2009): 10-18.
20. Wu, Xindong, et al. "Top 10 algorithms in data mining." Knowledge and Information Systems 14.1 (2008): 1-37.

21. Vītoliņa I., Kapenieks A. User analysis for e-inclusion in a blended learning course delivery context. Rēzeknes Augstskolas starpt konference "Sabiedrība, integrācija, izglītība", 2014. gada 23.-24. maijs
22. Vītoliņa I., Kapenieks A.. 2013. E-inclusion measurment by e-learning course delivery. "ICTE in Regional Development 2011-2012" Valmiera.
23. Vītoliņa I., Kapenieks A. 2013. E-inclusion and knowledge flows in e-course delivery. Proceedings of the 5th International Confere Computer Supported Education CSEDU 2013, Aachen, Germany, 6-8 May, 2013, pp. 417-422. Aachen: SCITEPRESS, 2013. 9789898565532
24. Zhang, Y., et al. Withdrawal prediction using the blackboard learning management system through SOM. Software Engineering an Mining (SEDM), 2010 2nd International Conference on. IEEE, 2010.
25. Nissen, M.,E. 2006. Harnessing knowledge dynamics: Principled organizational knowing & learning . p. 278.
26. Kongsakun, K. "An improved recommendation model using linear regression and clustering for a private university in Thailand." Ma Learning and Cybernetics (ICMLC), 2013 International Conference on. Vol. 4. IEEE, 2013.
27. Moseley, L. G., & Mead, D. M. Predicting who will drop out of nursing courses: A machine learning exercise, Nurse Education Today, pp.469-475,2008
28. Xenos, M. "Prediction and assessment of student behaviour in open and distance education in computers using Ba networks,"Computers and Education, 43(4), pp.345-359,2004
29. Lavendelis, E. Extending the MASITS Methodology for General Purpose Agent Oriented Software Engineering. In: *Proceedings of International Conference on Agents and Artificial Intelligence (ICAART)*, Portugal, Lisbon, 10-12 January, 2015. Lisbon: 2015, pp.15 ISBN 978-989-758-073-4.
30. Lavendelis, E., Grundspeņķis, J. MASITS – A Multi-Agent Based Intelligent Tutoring System Development Methodology. In: *Proce of IADIS International Conference „Intelligent Systems and Agents 2009"*, Portugal, Algarve, 21-23 June, 2009. Algarve: IADIS press pp.116-124. ISBN 978-972-8924-87-4.
31. Odell, James, H. Van Dyke Parunak, and Bernhard Bauer. "Extending UML for agents." Ann Arbor 1001 (2000): 48103.